~~SECRET~~
SECURITY INFORMATION ~~CONFIDENTIAL~~

69

9/G

7679

PROVISIONAL INTELLIGENCE REPORT

WEATHER-CROP YIELD CORRELATIONS AS APPLIED TO CROP YIELD ESTIMATES FOR THE EUROPEAN USSR

CIA/RR PR-15

26 May 1952

CENTRAL INTELLIGENCE AGENCY

OFFICE OF RESEARCH AND REPORTS

~~CONFIDENTIAL~~

~~SECRET~~

WARNING

This material contains information affecting the National Defense of the United States within the meaning of the espionage laws, Title 18, USC, Secs. 793 and 794, the transmission or revelation of which in any manner to an unauthorized person is prohibited by law.

SECRET
SECURITY INFORMATION

CONFIDENTIAL

PROVISIONAL INTELLIGENCE REPORT

WEATHER-CROP YIELD CORRELATIONS AS APPLIED TO CROP YIELD ESTIMATES FOR THE EUROPEAN USSR

CIA/RR PR-15

Note

The data and conclusions contained in this report do not necessarily represent the final position of ORR and should be regarded as provisional only and subject to revision. Additional data or comments which may be available to the user are solicited.

CENTRAL INTELLIGENCE AGENCY

Office of Research and Reports

CONFIDENTIAL

SECRET

CONTENTS

WEATHER-CROP YIELD CORRELATIONS AS APPLIED TO CROP YIELD ESTIMATES FOR THE EUROPEAN USSR

Summary

In this report a study is made of the relationship between the yields of three spring grains (wheat, barley, and oats) and selected weather factors in the major grain-surplus-producing districts of the European USSR. Total monthly precipitation and mean monthly maximum temperature values were selected as the weather factors to be employed. Precipitation values were used both in the form of separate monthly totals and in the form of combinations of monthly data. Only individual mean monthly maximum values were used for temperature. These weather data, as well as the grain yield information, cover the period 1883-1915.

From an analysis involving these factors, multiple regression equations have been developed in this study which, on the basis of current weather information, will be employed for the first time by ORR in estimating the yields of the 1952 grain crop in the USSR. The weather components employed in the development of these prediction equations were selected on the basis of comparative significance as determined through the use of simple and multiple correlation coefficients. The significance of certain of these correlation coefficients formed the basis for hypothesizing certain months or combinations of months as most "critical," with respect to precipitation and maximum temperature, in determining the ultimate crop

yield. The crop yield data were utilized in such a way that part of them was employed in obtaining preliminary information and in setting up the hypothesis as to "critical" months, and another part was used to test the validity of this hypothesis. The correlation coefficients computed in this study formed the basis for setting up the multiple regression equations to be used in estimating future grain yields in the USSR. However, since only a limited number of weather factors can be included in a prediction equation, a result obtained in any particular case may have to be readjusted in the light of any significant meteorological or nonmeteorological factor not considered in the equation.

On the basis of results obtained thus far, a continuation of this weather-crop yield study is planned. In particular, the relationship between crop yields and vapor pressure deficit values will be investigated. Vapor pressure deficit, as a function of both temperature and humidity, is a rough measure of the rate of transpiration and evaporation from plants. Preliminary investigations indicate that significant results may be obtained, particularly in regions of marginal precipitation.

In addition, studies on the major winter grains and other crops, such as potatoes, sugar beets, and cotton, are planned.

I. Purpose and General Introduction.

This report deals with a study of the problem of determining past weather-crop yield correlations and ascertaining whether it is practical to use these correlations in estimating future grain yields in the European USSR. This study is a progress report, and conclusions are based on analyses made up to the present time.

The importance of weather factors in relation to crop yields has long been recognized, but it has been extremely difficult to express these weather-crop yield relationships in mathematical terms that could be applied to estimating grain yields. Many separate weather factors affect the final crop yields, and the problem is further complicated by the fact that there are numerous interactions among the weather factors themselves. For example, in order to determine the effect of precipitation on the yield of spring wheat in any given area, not only the amount of precipitation but also the period of its occurrence must first be considered. In addition to precipitation, it is necessary to consider the effects of other weather factors occurring during the same period. The effects of these other factors are not independent of the precipitation or of each other but are interrelated. Thus the problem is exceedingly complex.

If practically unlimited time and data were available, it would be ideally desirable to measure the relationships between weather factors and crop yields by evaluating both the qualitative and the

quantitative effects of every conceivable combination of weather factors at regular intervals throughout the growing season. However, since both time and data are insufficient for such an idealistic approach, the analysis of only a limited number of weather factors is included in this study.

Since relatively comprehensive weather and crop yield information in the European USSR is available only for the period 1883-1915, it was necessary to make use of this information even though it is old. For the purposes of a correlation study, however, such data are entirely satisfactory, since the general pattern of the relationships involved does not change.

II. Area Covered.

The area selected for this study includes most of the grain-surplus-producing districts of the European USSR.* With its apex based in the Tsarist guberniya** of the Don the selected area spreads away fanwise:

1. Northwestward, including the following guberniyas of the Ukraine: Yekaterinoslav, Poltava, Chernigov, and Kiev as well as the southwest guberniya of Bessarabia;

2. Northward, including the following guberniyas of the "Black Soil Belt": Voronezh, Kursk, Orël, Tambov, and Penza as well as the semi-Black Soil guberniyas of Tula and Ryazan;

3. Northeastward, including the Volga River valley guberniyas of Saratov, Samara, Simbirsk, and Kazan as well as the Ural guberniya of Perm.

In terms of present-day (1952) boundaries, all or at least portions of all of the following administrative divisions are located in this area with its apex based in Rostov oblast of the former Don guberniya (the northwestern part of Economic Region IV):

1. To the northwest in the Ukrainian group (Economic Region III) are located the following: Stalino, Voroshilovgrad,

---

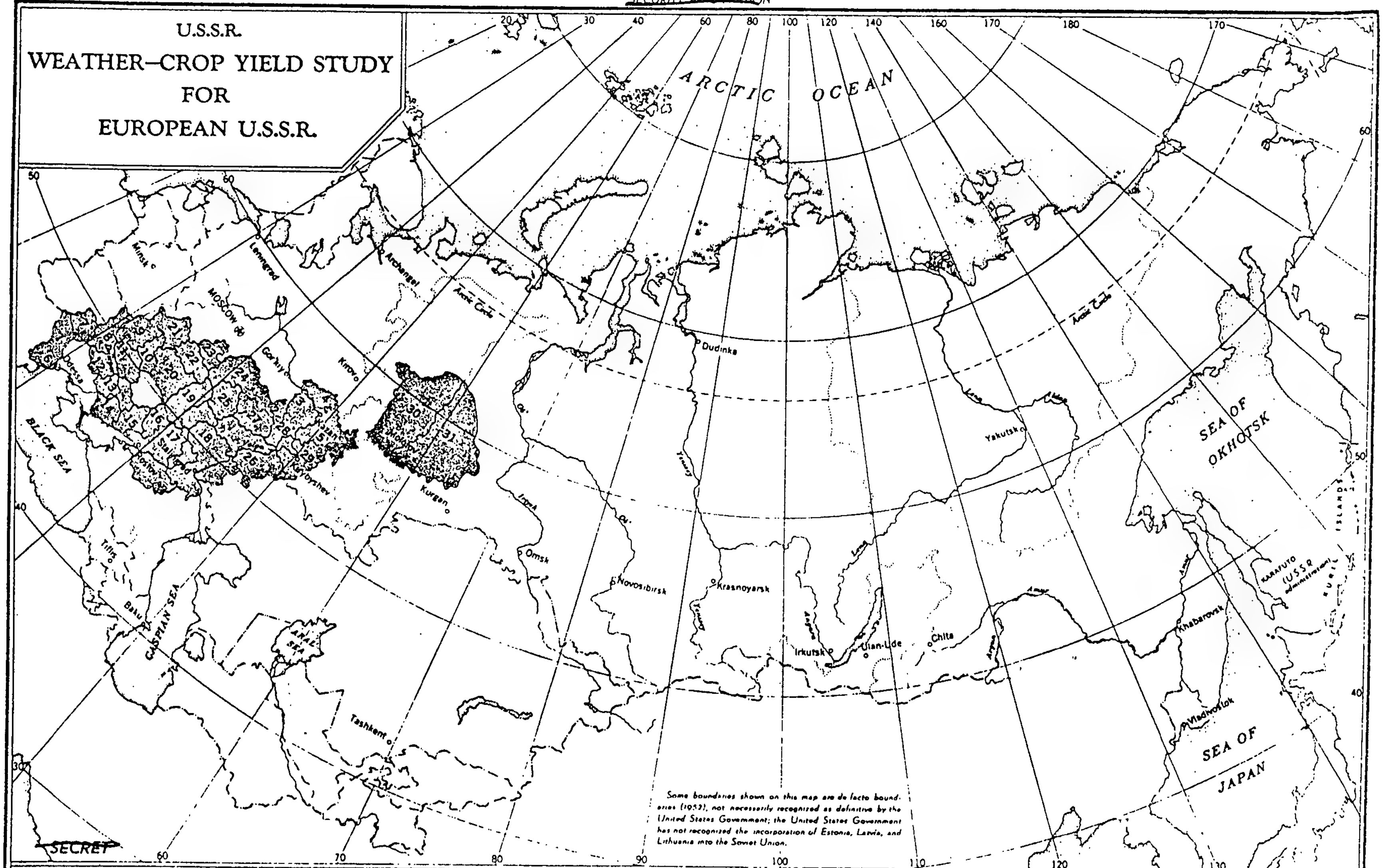
* This area is outlined in the map following p. 6.
** Guberniyas, the larger administrative divisions in the early 1900's, correspond roughly to present-day oblasts, although in many cases they were somewhat larger. The next smaller administrative divisions in this period, the uyezds, though larger, are roughly comparable to the present-day rayons.

Zaporozh'ye, Dnepropetrovsk, Kirovograd, Poltava, Sumy, Chernigov, Kiev, and Vinnitsa oblasts and the Moldavian SSR and Izmail oblast;

2. To the north in the Black Soil and semi-Black Soil Belts (the southern part of Economic Region VII) are located the following: Kursk, Orel, Tula, Bryansk, Voronezh, Tambov, Ryazan, and Penza oblasts and Mordovskaya ASSR, Chuvashskaya ASSR, and Mariyskaya ASSR;

3. To the northeast in the Volga group (Economic Region VI) are located the following: Stalingrad, Saratov, Kuybyshev, and Ul'yanovsk oblasts and Tatarskaya ASSR. In addition, two Ural (Economic Region VIII) oblasts are included: Molotov west of the Urals and Sverdlovsk east of the mountains.

SECURITY INFORMATION



International Boundary
Union Republic (S.S.R.)
Autonomous Republic (ASSR)
Oblast (selected)

0 250 500 750 Statute Miles
0 250 500 750 Kilometers

AREAS INCLUDED IN STUDY

UNION REPUBLIC
1. Moldavian S.S.R.

AUTONOMOUS REPUBLICS (In R.S.F.S.R.)
2. Mordovskaya ASSR
3. Chuvashskaya ASSR
4. Mariyskaya ASSR
5. Tatarskaya ASSR

OBLASTS (selected) IN UKRAINIAN S.S.R.
6. Izmail
7. Vinnitsa
8. Kiev
9. Chernigov
10. Sumy
11. Poltava
12. Kirovograd
13. Dnepropetrovsk
14. Zaporozh'ye
15. Stalino
16. Voroshilovgrad

OBLASTS (selected) IN R.S.F.S.R.
17. Rostov
18. Stalingrad
19. Voronezh
20. Kursk
21. Bryansk
22. Orël
23. Tula
24. Ryazan'
25. Tambov
26. Saratov
27. Penza
28. Ul'yanovsk
29. Kuybyshev
30. Molotov
31. Sverdlovsk

12223 CIA, 2-52

III. Sources of Information and Its Tabulation.

The raw materials for this study were obtained from two sets of Russian data. The crop yield data were taken from Urozhay (Annual Publications of the Central Statistical Committee). 1/ These volumes contain acreage and yield data for the major crops in Russia for each year during the period 1883-1915. The yields and acreages are given on both an uyezd and a guberniya basis.

The weather information was obtained from Letopici (Annals of the Central Physical Observatory). 2/ The weather data in those volumes issued during the interval 1883-1915 are comparable with the selected yield data. Rather complete weather information on a monthly, and in some cases even on a daily, basis for individual stations is contained in each of these yearly volumes. Despite the large number of stations tabulated in each yearly volume, it was difficult, in some regions, to find stations which had a long-term record. The reason for this is that the same stations were not always included in every volume. Sometimes old stations would be omitted and new ones added.

From the set of weather volumes, total monthly precipitation and mean monthly maximum temperature values were tabulated for all stations in the guberniyas listed above. Only stations having at least 15 years of record within the period 1883-1915 were selected. The location of each of these stations was then determined both as

to guberniya and as to uyezd in order that yield tabulations might be made for each guberniya and for those individual uyezds in which weather stations were located.

Tabulations were then made of the yields for three spring grains (wheat, barley, and oats) for each individual uyezd which contained at least one weather station having 15 years or more record and for each guberniya in which such uyezd was located. From 1883-1894, yields were recorded in chetverts per dessiatine, while from 1895-1915 they were expressed in poods per dessiatine. For purposes of conformity, the chetvert values were converted to the equivalent poods.*

---

* The following conversion factors are applicable for wheat:

1 chetvert per dessiatine (2.70 acres)	= 9.903 poods per dessiatine
1 chetvert per dessiatine	= 1.485 centners per hectare (2.47 acres)
1 pood per dessiatine	= 0.150 centner per hectare
1 chetvert	= 5.96 bushels (357.60 pounds or 162.21 kilograms)
1 pood	= 0.60 bushel (36.11 pounds or 16.38 kilograms)
1 centner	= 3.67 bushels (220.46 pounds or 100 kilograms)

IV. Purpose and Method of Eliminating the Yield Trend.

Since the yield data extended over a relatively long period of years, it was necessary to perform preliminary tests to determine whether a yield trend existed for which an allowance should be made. These tests, performed in the manner described below, indicated the presence of an increasing trend in yields over the period 1883-1915. This indicated trend was found to be in agreement with a statement by Timoshenko 3/ that "the average yield of all grains increased during the period 1883-1915 on the average by half a pood per dessiatine yearly, or about 1 percent of the average yield for the period." The method of "least squares" was used for the determination of the linear trend. Under the assumption of a linear trend the formula will be of the type $Y = a + bX$.* The "b" value in the formula indicates the slope of the trend line. For example, a value of $b = +1.0$ would indicate that the yield increased 1 pood per dessiatine each year and that such a yield increase was not caused by weather factors but rather by increasingly improved technology, better varieties, or other similar factors. Since this is a study on the relationship between weather and crop

* This formula, however, describes any one of an infinite number of lines, and the problem is to determine which line best describes the data. The principle of "least squares" is used in determining this best line. The line of best fit to each series of yields is a line about which the sum of squares of the deviations (the differences between the line and the actual yields) will be a minimum. There can be only one such line.

yields, it was necessary to eliminate this variation, which was not caused by weather factors. Hence all yield data used in this study have been adjusted for this trend.

The possibility of a significant yield trend from 1915 to the present also poses a serious problem in arriving at current production estimates. The question is continually raised as to whether such factors as lime and fertilizers, increased mechanization, improved farm practices, and improved varieties have significantly increased yields over a period of years up to the present time. Concerning fertilizers, Jasny 4/ states that "commercial fertilizers have been applied to only a few crops, mainly sugar beets, cotton, and, to a smaller extent, flax. Such crops as grain benefited only in so far as they were grown in rotation with fertilized crops." Likewise there is no evidence available that mechanization of agriculture has had an appreciable effect in increasing grain yields. As to improved varieties, an external research project is being tentatively outlined in an effort to determine what effect, if any, they may have had in increasing yields.

The intelligent use of any yield prediction equations subsequently developed in this study requires a consideration of any nonmeteorological factors that can be shown to have had a significant effect in increasing yields.

V. Methodology and Results.

1. Weather Factors Used.

Having tabulated the weather and yield data and corrected the yields on the basis of the indicated yield trend, the next step was a joint analysis of the data in the three series of tabulations -- total monthly precipitation, mean monthly maximum temperature, and yields of the selected crops. Before proceeding with a discussion of the methodology employed, however, it would perhaps be beneficial to give an explanation and justification for the use of precipitation and maximum temperature as the weather factors to be used in this study. Precipitation has been one of the factors employed in most studies of weather-crop yield relationships. This is not surprising in view of the fact that moisture is an essential factor in all crop-producing areas. In regions of marginal or submarginal precipitation -- for example, the valleys of the lower Volga and the Don as well as parts of South Ukraine -- moisture is the all-important factor. In these regions an average or even above-average rainfall is necessary for even mediocre crop production. Furthermore, the distribution of the rainfall throughout the year, and particularly through the growing season, is of the utmost importance. It is generally recognized that, during certain phases of development, plants are more sensitive to environmental conditions and are more easily damaged by extremes. Some authorities have termed these periods of stress as "critical" stages in the

growth of plants. For example, numerous scientists in the US, Canada, and Russia have considered the period just prior to the heading of wheat as a "critical" stage in determining the ultimate crop yield. Adequate moisture must be available during this "critical" period if a "good" yield is to be obtained.

The reason for the use of maximum temperatures is perhaps less obvious. As previously mentioned, in many regions it was difficult to find stations having a long meteorological record. In the studies employing uyezd yields, precipitation and temperature data were, as a rule, available from only one station. Even when guberniya yields were involved, weather data from only a limited number of stations were available, varying from three to eight or nine. An average of the precipitation at this number of stations within a guberniya is not necessarily representative of the guberniya as a whole. This is particularly true during the growing season, since much of the precipitation during the summer occurs in convectional thunderstorms, which are characteristically local in nature although frequently quite intense. Temperatures, on the other hand, are much less variable over a particular area than is precipitation. Furthermore, average maximum temperatures are quite closely correlated with actual precipitation. General cloudiness, often occurring with local thunderstorms over a particular area, will be reflected in lower maximum temperatures. For

this reason, the less variable average maximum temperature might give a better clue to actual precipitation over a large area than would the average precipitation values recorded at a limited number of stations.*

Employing the yield data and monthly data for the two weather factors, precipitation and maximum temperature, methods of correlation and regression were employed to determine the mutual relationships among these data first on an uyezd basis and then on a guberniya basis.

2. Simple Correlations.

Simple correlation coefficients** were computed between the uyezd yields of each of three spring grains (wheat, barley, and oats) and monthly precipitation or a combination of the recorded data for such combinations of months as might possibly have some bearing on yield. Similar coefficients also were computed between yield data and temperature values.

---

* It is also conceivable that since maximum temperature is quite sensitive to cloudiness, there may be some relationship between maximum temperature and the rate of evaporation and transpiration by plants.

** Since yields are directly dependent on weather factors, use of the term <u>regression coefficient</u> rather than <u>correlation coefficient</u> would be technically more nearly correct. The distinction between the two is that regression coefficient is the appropriate term if the one variate, in this case yields, may be considered as dependent upon the other, in this case, precipitation or maximum temperature. Correlation, on the other hand, is the appropriate measure of the relation between two variates neither of which may be looked upon as a consequence of the other. Correlation has acquired many of the concepts of regression, and often the distinction between the two has been almost lost.

Even a superficial examination of certain coefficients computed with the use of monthly data or a combination of monthly data indicated that in certain cases little or no relationship existed. In other cases a more or less positive pattern of relationships was indicated. Based on these indicated relationships, a working hypothesis was assumed as to the "critical" months with respect to both precipitation and temperature for each uyezd under consideration.

The next step was to test, on the basis of a guberniya as a whole, the working hypotheses assumed for the several uyezds within that guberniya. For these tests the method of simple correlations was employed.

For example, the simple correlation coefficients between spring wheat yields in Samara uyezd of Samara guberniya* and the specified weather factors at Polibino, a station in Samara uyezd, are as follows:

---

* Computations for Samara guberniya are used as examples throughout this study. The boundaries of the Tsarist guberniya include all or at least portions of the following present-day oblasts: Stalingrad, Saratov, Kuybyshev, and Ul'yanovsk.

Simple Correlation Coefficients for Spring Wheat
in Samara Uyezd of Samara Guberniya

Weather Factor	Simple Correlation Coefficient, r a/
Precipitation	
Sep-Oct plus May-Jun	.547** b/
Feb-Jun	.529**
Mar-Jun	.537**
Apr-Jun	.582**
May-Jun	.647**
Average Maximum Temperature	
May	-.370
Jun	-.722**
Jul	-.516**

a. A correlation coefficient, generally designated as r, is an attempt to summarize in one number the degree of relationship existing between two series of observations: for example, wheat yields and average June precipitation. Or it might be defined as the degree to which the wheat yields and precipitation values keep in step as they increase or decrease in amount. The correlation coefficient is designed to vary between -1 and +1. The two extremes indicate perfect linear relationship. A negative value merely indicates an inverse relationship: in other words, an increase in one variable is associated with a decrease in the other variable. A zero value of r indicates that the two variables are independent.

b. Double asterisks in tables indicate a highly significant correlation: that is, significant at the 1-percent level. In other words, the chances are only 1 out of 100 of obtaining a correlation as large as this due to chance alone. The levels of significance were obtained by use of the tables in Statistical Methods by George W. Snedecor. 5/

Twenty-seven observations were used in computing each of the above coefficients. All the coefficients except that for May average maximum temperature are highly significant: that is, significant at the 1-percent level.

Other simple correlation coefficients for various uyezds are given in Appendix A, Tables 1, 2, and 3. A superficial examination of the data in these tables suggests that on an uyezd basis May-June appear to be more nearly "critical" months as regards precipitation in most guberniyas, while June appears more nearly "critical" as regards maximum temperature. The validity of this apparent significance will be tested later in this study.

It is obvious that there should be variations among these correlations on an uyezd level, and it is difficult to attach any great significance to individual values within a given uyezd. The primary purpose of computing these coefficients on an uyezd level was to gain preliminary information as to the nature of the variability among various localities as well as to obtain a rough idea as to which weather factors appeared to play a predominant role within these various localities.

The correlation coefficients based on the individual uyezd yields formed the basis for setting up a working hypothesis to be tested on the basis of guberniya yields. Main emphasis was placed on a study of the three spring grains previously mentioned.

As previously stated, numerous correlation studies in the US, Canada, and Russia have indicated that the period just prior to the heading of grains is "critical" in determining the ultimate crop yield. The heading stage for spring grains in most of the area studied occurs within the period of late May to early July. For this reason, the sum of the May and June precipitation was used as one portion of the working hypothesis to be tested in individual guberniyas by the simple correlation method. It was felt that the sum for the 2 months would be a more realistic value to use than the precipitation for any single month.

An attempt was made, using uyezd correlation coefficients for temperature, to determine for each guberniya which months appeared to be most "critical." For example, if the correlation coefficients for several uyezds within a certain guberniya seemed to indicate that June temperature was most closely correlated with yield, then, in that guberniya, the June temperature factor was used as the second portion of the working hypothesis to be tested.

In Samara guberniya, for example, the average maximum June temperature and the sum of the May and June precipitation were selected as the two portions of the hypothesis to be tested.

Under this hypothesis the correlations between spring wheat yields in Samara guberniya and the hypothesized weather factors are as follows:

Simple Correlation Coefficients for Spring Wheat in Samara Guberniya

Weather Factor	Simple Correlation Coefficient, r
Sum of May and June Precipitation	.719** a/
Average Maximum June Temperature	-.560**

a. Double asterisks in tables indicate a highly significant correlation: that is, significant at the 1-percent level.

The above coefficients are based on 28 observations, and both are highly significant.

These correlations and others similarly computed for various guberniyas are shown in Appendix A, Table 4. Simple correlation coefficients for precipitation and temperature using guberniya yields have been computed for 15 of the 18 guberniyas listed above. Correlation coefficients for temperature on an uyezd basis indicated that in 9 of the 15 guberniyas the average maximum temperature for June appeared to be more "critical" than the temperature for May or July. Therefore, in the case of each of these nine guberniyas, the June temperature was selected as the temperature

portion of the hypothesis. In the case of a central block of six guberniyas (Tambov, Ryazan, Tula, Orel, Kursk, and Voronezh), similar coefficients indicated the average maximum temperature for May as being more "critical" than for June or July, and in the case of each of these six guberniyas, therefore, May temperature was selected as the temperature portion of the hypothesis. In certain of the 15 guberniyas, correlations could not be made on an uyezd basis. In such cases the hypothesis was formulated using the results of uyezd correlations in neighboring guberniyas.

A superficial examination of the data in Table 4 indicates that in the case of each of three spring grains (wheat, barley, and oats) there frequently is a closer correlation between temperature and yield than between precipitation and yield. Also, the correlation coefficients for precipitation are significant in more cases for barley and oats than for wheat.

As a rough check on the accuracy of the working hypothesis previously assumed, simple correlations between spring grain yields and temperatures were computed for the months adjacent to the month of June, which had been set up in the hypothesis as being "critical." For example, in Samara guberniya the correlations between spring wheat yields and average maximum May and July temperatures are as follows:

Simple Correlation Coefficients for Spring Wheat in Samara Guberniya

Weather Factor	Simple Correlation Coefficient, r
Average Maximum May Temperature	-.412* a/
Average Maximum July Temperature	-.530**

a. A single asterisk in tables indicates a significant correlation: that is, significant at the 5-percent level. Double asterisks in tables indicate a highly significant correlation: that is, significant at the 1-percent level.

These two coefficients, based on 28 observations, are both lower than that for the hypothesized "critical" month of June, which was -.560. These and other correlations computed on a similar basis are shown in Appendix A, Table 5.

An examination and comparison of the temperature correlation coefficients in the six guberniyas in which the May average maximum temperature was hypothesized as "critical" raises some question as to the validity of this assumption.

Considering the 18 sets of correlations (each set of three being composed of the temperature correlations for May, June, and July) involving three spring grains (wheat, barley, and oats) in each of the six guberniyas, the June temperature correlations were highest in nine sets, May values were highest in five sets, and July highest in four. In some instances, however, the differences between the

correlations for particular months could not be considered significant. For oats, the June temperature correlations in all six guberniyas were higher than those for May. In view of these results, it might have been better to have hypothesized June temperature as "critical" for all guberniyas in this study rather than to have made an exception for the six guberniyas discussed above.

Considering the similar sets of correlations in the other 9 guberniyas where June temperature was hypothesized as critical, there were only 6 cases out of the 27 in which either the May or July temperature correlation coefficients were significantly better than the hypothesized June coefficients. The correlation coefficients in two guberniyas, Bessarabia and Yekaterinoslav, were consistently small, in most cases not even significant at the 5-percent level. If the coefficients for these two guberniyas are omitted from the tabulation, in only 2 of the remaining 21 sets of correlations was either the May or the July coefficient significantly higher than that for the hypothesized month of June.

3. Multiple Correlations.

As pointed out earlier in this discussion, the fluctuations in a given yield series are never dependent upon one single weather factor. The next step, therefore, was to combine the two selected weather factors into a multiple correlation with yield. Multiple correlation consists of the measurement of the relationship between

a dependent variable, in this case yield, and two or more independent variables, in this case monthly precipitation and maximum temperature values.

To compute a multiple correlation coefficient for three variables Y, X_1, and X_2, it is first necessary to compute three simple correlation coefficients r_{Y1}, r_{Y2}, and r_{12}. The next step is to compute two standard partial regression coefficients, $b'_{Y1.2}$ and $b'_{Y2.1}$, from the following relationships involving the simple correlations

$$b'_{Y1.2} = \frac{r_{Y1} - r_{Y2}\, r_{12}}{1 - r_{12}^2} \quad \text{and} \quad b'_{Y2.1} = \frac{r_{Y2} - r_{Y1}\, r_{12}}{1 - r_{12}^2}$$

Then the square of the multiple correlation coefficient is obtained by

$$R^2 = r_{Y1}\, b'_{Y1.2} + r_{Y2}\, b'_{Y2.1}$$

The square root of this value gives us the desired multiple correlation coefficient, R. The multiple correlation coefficient, R, is always less than unity but is greater than any of the simple correlation coefficients that enter into its computation.

Multiple correlation coefficients were first computed by using only the weather factors for those months assumed under the working hypothesis to be "critical." For example, for Samara guberniya the components were as follows: the yields of the three

spring grains (wheat, barley, and oats), the sum of May and June precipitation, and average maximum June temperature. As shown in Appendix A, Table 6, the multiple correlation coefficient for spring wheat in Samara guberniya based on 28 observations is .753, a highly significant value. Of the 39 individual multiple correlation coefficients, computed under similar hypotheses and given in Table 6, 27 are highly significant (1-percent level of significance) and 7 other values are significant (5-percent level of significance).

As a check on the multiple correlation coefficients computed under the working hypotheses, other coefficients were computed by using temperatures for months adjacent to the one considered "critical." For example, in the case of Samara guberniya, as indicated in Appendix A, Table 7, the multiple correlation coefficient based on spring wheat yield, the sum of May-June precipitation, and average maximum May temperature is .734. Similarly, the coefficient based on spring wheat yield, the sum of May-June precipitation, and average maximum July temperature is .729. Both values are highly significant but slightly less than the coefficient, .753, computed under the hypothesis based on June average maximum temperature. In Table 7 will be found other multiple correlation coefficients for spring grains computed by using temperatures for months adjacent to the one considered "critical."

In the discussion of simple correlations the question was raised as to the validity of hypothesizing May average maximum temperature as "critical" in six guberniyas* rather than June temperature, which was hypothesized for the remaining guberniyas. A close examination of the multiple correlation coefficients in Appendix A, Table 7, casts further doubt on the validity of hypothesizing either May or July average maximum temperature as "critical." Considering the 18 sets of multiple correlations (each set of 3 being differentiated as to May, June, or July temperature) involving the 3 spring grains (wheat, barley, and oats) in each of the 6 guberniyas*, the correlations based on June temperature were highest in 9 sets, while May values were highest in 5 sets and July highest in 4. These results are identical with those obtained with simple correlations. On this basis, therefore, it seems desirable to reject that portion of the temperature hypothesis wherein the May average maximum temperature was considered "critical" for six guberniyas and, in its place, to consider June temperature as "critical" for all guberniyas.

Considering the similar sets of correlations in the other 7 guberniyas where June temperature actually was hypothesized as "critical," there were only 6 cases out of the 21 in which the coefficients based on May or July temperatures were higher than those based on June temperatures. Furthermore, in several of these six cases the differences in the correlations were slight.

* Tambov, Ryazan, Tula, Orël, Kursk, and Voronezh.

One other attribute of the given multiple correlation coefficients might shed some light on their interpretation. If the value of R, the multiple correlation coefficient, is squared, the result is a measure of the proportion of the variability in yield that can be attributed to the precipitation and temperature values used in computing R. For example, assume that a multiple correlation coefficient R = 0.80 is obtained by using the sum of May and June precipitation, the average maximum June temperature, and spring wheat yields in a given guberniya. It can then be stated that the square of 0.80, or 64 percent of the year-to-year variability in the spring wheat yields, has been accounted for by the two given weather factors and that only 36 percent remains unaccounted for.

4. Multiple Regression Equations for Forecasting Yields.

The final step to be taken to obtain a formula suitable for use in forecasting yields in a given area is to employ multiple correlation coefficients, computed under a given hypothesis, to develop multiple linear regression equations for two independent

variables of the type $\hat{Y} = a + b_{Y1.2} X_1 + b_{Y2.1} X_2$.* In this equation, $\hat{Y}$ is the estimated or predicted yield, and X_1 and X_2 are the independent variables, precipitation and temperature.

The expressions $b_{Y1.2}$ and $b_{Y2.1}$ are partial regression coefficients. The first expression, $b_{Y1.2}$, can be read as "the regression of Y on X_1 independent of X_2," while $b_{Y2.1}$ can be read as "the regression of Y on X_2 independent of X_1." Assuming that Y = spring wheat yield in poods, that X_1 = the sum of May and June precipitation in millimeters, and that X_2 = the average maximum June temperature in degrees centigrade, then the expression $b_{Y1.2}$ indicates the quantity of change (in terms of poods) in the yield of wheat for each millimeter change in the amount of May and June

---

* This formula is a mathematically simplified form of the following equation:

$$\hat{Y} = \bar{y} + b'_{Y1.2} \frac{\sqrt{\sum y^2}}{\sqrt{\sum x_1^2}} (X_1 - \bar{x}_1) + b'_{Y2.1} \frac{\sqrt{\sum y^2}}{\sqrt{\sum x_2^2}} (X_2 - \bar{x}_2)$$

This longer form of the equation has an advantage in that the component parts of the regression equation can be more easily identified. Here $\hat{Y}$ is the estimated or predicted Y; $\bar{y}$ is the mean of all the given Y values; $b'_{Y1.2}$ and $b'_{Y2.1}$ are the two standard partial regression coefficients described above; X_1 and X_2 are the independent variables, while $\bar{x}_1$ and $\bar{x}_2$ are their respective means; $\sum y^2$, $\sum x_1^2$, and $\sum x_2^2$ are the sums of squares of deviations from the mean of Y, X_1, and X_2, respectively.

precipitation with the actual average maximum June temperature remaining constant. Similarly, the expression $b_{Y2.1}$ indicates the quantity of change (in terms of poods) in the yield of wheat for each degree centigrade change in the average maximum June temperature with the actual May-June precipitation remaining constant. In other words, the coefficients indicate the net relationship between the dependent variable, yield, and one of the independent variables (for example, the sum of May and June precipitation) while allowing for the other independent variable (for example, average maximum June temperature), which also is considered in computing the coefficient.

In the equation $\hat{Y} = a + b_{Y1.2} X_1 + b_{Y2.1} X_2$ the values a, $b_{Y1.2}$, and $b_{Y2.1}$ are all constants that can be quite readily computed from the given data on yields, precipitation, and temperature. For example, in computing the multiple regression equation for use in predicting spring wheat yields in Samara guberniya, actual yields, May-June precipitation, and average maximum June temperature for 28 years were taken into consideration to obtain the values a, $b_{Y1.2}$, and $b_{Y2.1}$ which resulted in the equation $\hat{Y} = 55.9 + 0.325X_1 - 2.049X_2$ as given in Appendix A, Table 8. Table 8 also contains other multiple regression equations, based on May-June precipitation and average maximum June temperature, for use as an aid in predicting spring grain yields in selected guberniyas of the USSR. If the sum of May and June precipitation

in any particular year is substituted for X_1 and the average maximum June temperature is substituted for X_2, it is possible to compute the estimated or predicted yield, $\hat{Y}$, for that year. Some of these prediction equations are more reliable than others. As a general rule, those guberniyas for which relatively high multiple correlation coefficients were obtained for the various spring grains will tend to have more reliable prediction equations.

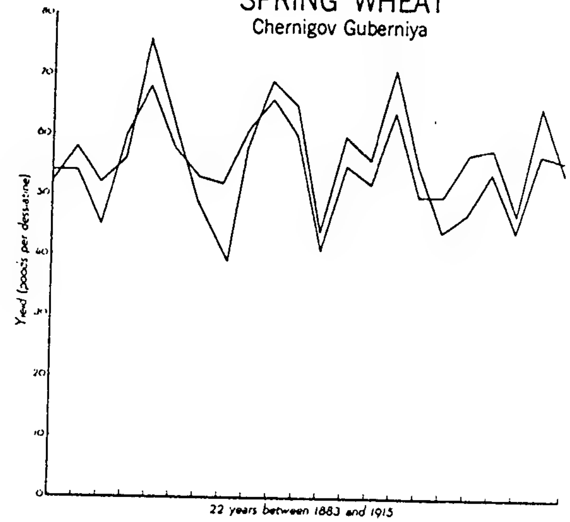
Some of the results obtained by using the multiple regression equations for computing yields of various grains in the interval 1883-1915 and comparing them with the actual known yields are shown graphically in Figures 1 to 7. Failures in accurately estimating the yields result in what are known as errors of estimate. The standard error of estimate, defined as $s_{Y.12} = \sqrt{\frac{(1-R^2)\sum y^2}{n-3}}$,* is a measure of the variation among errors of estimate. The greater the standard error, the poorer is the relationship between computed and actual yields. This standard error of estimate is useful in establishing limits as to the accuracy of the estimation. For example, if we assume a standard error of estimate of 5 poods, it is possible to state that of all the estimates based on the regression equation, roughly two-thirds of them will be correct to within plus or minus 5 poods. Stated

* In this equation, R is the multiple correlation coefficient; $\sum y^2$ is the sum of squares of deviations about the mean of Y; and n is the number of observations in the series, or the number of years of yield and weather records used in computing the multiple regression equation.

Figure 1

SPRING WHEAT

Chernigov Guberniya



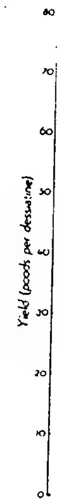
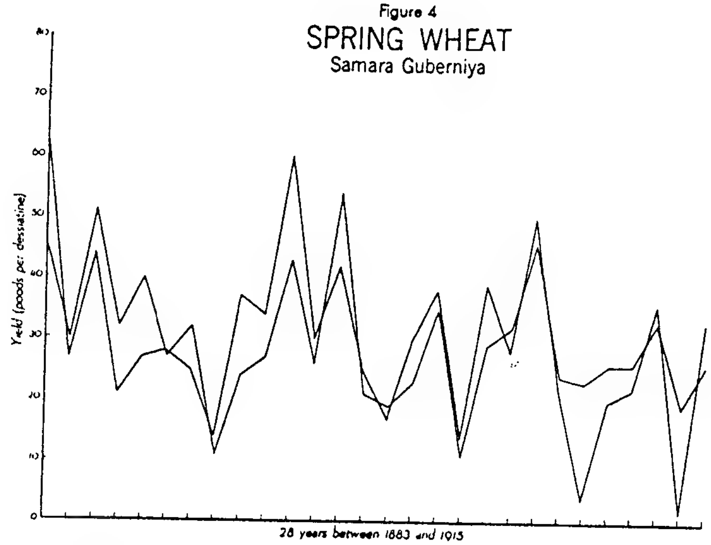
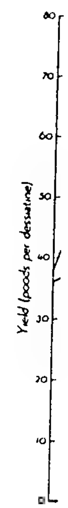


Figure 4

SPRING WHEAT

Samara Guberniya





GR1136 CIA, 4 52

another way, the chances would be 67 out of 100 that any one estimate would be correct to within plus or minus 5 poods. Similarly, the factor 1:33 times the standard error of estimate will include roughly 80 percent of the cases for the regression equations developed in this study.

For the graphs shown in Figures 1 to 7, the values of the multiple correlation coefficient, R, and the factor 1.33 times the standard error of estimate, $s_{Y.12}$, are given in the table below. The factor 1.33 $s_{Y.12}$ merely means that, using the given data, 80 percent of all estimates based on the regression equation would be correct to within plus or minus the indicated number of poods.

Multiple Correlation Coefficients and Limits in Errors of Estimation

Figure	R	1.33 $s_{Y.12}$
1	.760	8.5
2	.764	10.6
3	.815	7.4
4	.753	12.8
5	.591	10.3
6	.798	10.0
7	.701	12.5

For example, Figure 4 shows the actual yearly spring wheat yields in Samara guberniya versus the yields computed using the regression equation $\hat{Y} = 55.9 + 0.325X_1 - 2.049X_2$. The average yield for the 28 years used in the computations was 30 poods. The highest yield was 62 poods, and the lowest was 2 poods. Assuming that the relations studied remain essentially constant, roughly 80 percent

of all prediction estimates based on this regression equation will be correct to within plus or minus 12.8 poods.

5. Assumptions.

At this point it is desirable to state several assumptions made in this study and to discuss their validity briefly. The first assumption is that the relationship between yield and any weather factor is linear. The validity of this assumption appears to depend largely on the area being considered. For example, where wheat is grown under meteorologically optimum conditions, it would certainly not be safe to assume that the yield would increase indefinitely as the June rainfall increased. Up to a certain point an increase in precipitation would result in higher yields, but beyond this point further precipitation might even lower the ultimate yield. The assumption of linearity appears justifiable, however, as long as the weather factors considered do not fluctuate over too wide a range. For example, in most of the major wheat areas in the USSR it is extremely unlikely that precipitation amounts will be as great as to cause reduced yields. In fact, lack of moisture is far more likely to cause low yields. In such cases, any slight curvilinear relationship that does exist usually can be satisfactorily approximated by linear methods. Furthermore, when linear methods are applied to curvilinear data, the degree of relationship is really greater than that indicated by the correlation coefficient. It is possible to compute curvilinear regression equations, but they are more

complicated and should be used only with lengthy series of observations.

A second assumption made in this study is that the effects of the independent variables (precipitation and maximum temperature) on the dependent variable (yield) are additive -- that is, a given change in either precipitation or maximum temperature has the same effect on yield regardless of the size of the other independent variable. In the linear multiple regression equation discussed previously, $\hat{Y} = a + b_{Y1.2} X_1 + b_{Y2.1} X_2$, the effect of a given change in X_1 on the size of $\hat{Y}$ is constant regardless of the size of X_2. The effect of X_1 on $\hat{Y}$ is independent of X_2. In a joint relationship, on the other hand, the effect of X_1 on $\hat{Y}$ is dependent on X_2. The effect of X_1 on $\hat{Y}$ depends on the size of X_2. The computation of a linear joint regression equation is considerably more complicated. Several of these computations, however, were made as a rough check on the validity of assuming additive rather than joint relationships. The equation used was of the form $\hat{Y} = a + bX_1 + cX_2 + dX_1X_2$. In this equation the product of the two independent variables, X_1X_2, expresses the joint relationship. This method in the particular instances used did not significantly improve the accuracy of the estimated yields. It would be unsafe to assume categorically that no improvements could be made by using this or other combinations of joint relationships, but it is questionable whether the improvement would be sufficiently great to warrant the necessary increase

in the expenditure of man-hours.

It would have been possible in this study to include multiple regression equations based on more than three variables, but an increase in the number of variables results immediately in increasingly complex computations. Furthermore, unless a fairly large number of observations are available, a prediction equation involving a large number of variables may not present a true picture. The equation may accurately fit the data from which it was derived but when applied to other similar data may not give satisfactory results.

VI. Sources of Current Weather Data.

Current weather information on the USSR is being received from several sources. The Department of Military Climatology, Air Weather Service, USAF, Andrews Field, is cooperating closely in supplying detailed information, and, under proposed new arrangements, precipitation totals by 10-day intervals will be available during the crop season as well as mean maximum temperatures for the corresponding intervals. During the remainder of the year, precipitation amounts will be supplied as monthly totals. During the winter months, any available snow cover conditions will be reported, as well as minimum temperatures in areas with little snow cover -- that is, areas in which danger from winter kill is the greatest.

Translated excerpts from the Soviet newspaper Socialist Agriculture (Sotsialisticheskoye Zemledeliye) are obtained from the American Embassy in Moscow. These excerpts contain, among other items, daily weather information (temperature and a general description of areas and intensities of precipitation), crop stage reports (stages of growth for various grains by regions), occasional crop condition reports (for example, "Condition of spring grains in Upper Volga is good, with an above-average yield expected"), and oblast procurement reports stating whether the oblast has fulfilled its grain delivery plan. The weather data and crop stage reports are tabulated on a daily basis by regions for most efficient use. Frequently, sufficient information concerning the stage of development

of a particular grain -- for example, the heading stage of spring wheat -- is obtained, so that it is possible to chart the northern movement of the stage. Such information is extremely valuable when used in conjunction with current weather information.

FDD also furnishes current weather and crop information obtained from the Soviet newspaper Sovkhoznaya Gazeta and other provincial newspapers. This information is included in the tabulations described above. Finally, the weekly FBIS Abstracts contain weather survey sections giving general temperature and precipitation data tabulated on a daily basis by regions -- useful as corroborative and supplementary information.

VII. Additional Proposed Investigations.

Further investigations of the relationships between weather factors and yields in the area covered by this study and in other areas of the USSR are contemplated. In particular, the effect of vapor pressure deficit on crop yields will be examined. Vapor pressure deficit is a function of the amount of moisture in the air and the temperature of the air and is technically defined as the difference between the actual vapor pressure of the atmosphere and the vapor pressure of a saturated atmosphere at the same temperature. As functions of temperature the vapor pressure deficit values are likely to show less variability over a region than does the average of the precipitation figures from a limited number of stations. Also, as functions of humidity, the vapor pressure deficits are a measure of the rate of transpiration and evaporation from plants. Significant correlation coefficients may be obtained, particularly in a region of marginal precipitation.

Further and more intensive studies on the major grains are planned, particularly the winter grains. In addition, other crops, such as potatoes, sugar beets, and cotton, will be investigated.

At the present time a list of meteorological stations in the USSR having a record of monthly precipitation values over a period of at least 20 years is being compiled. All known sources, including many original Russian language publications, are being used in an

effort to obtain the maximum length of record for each station. The use of such long-term records makes possible considerably more accurate statements as to whether a particular crop area in a given year has monthly precipitation amounts which are greater or less than the normal.

After a list of stations with a satisfactory length of record has been compiled, the monthly precipitation values will be tabulated. The stations will be broadly grouped according to primary agricultural regions, such as Lower Volga, Urals, Central Black Soil Zone, and then within each region they will be grouped according to oblasts.

Such a complete set of comparatively reliable mean values will be extremely helpful during the growing season as an indication of whether current precipitation amounts, as compared with the mean values, point to a potentially above- or below-average crop yield.

All the items discussed in this study have one ultimate goal -- the attainment of improved accuracy in the estimation of Soviet crop yields. One of the first steps toward this goal is the derivation of some of the underlying relationships between crop yields and the numerous interrelated weather factors. If some apparently significant relationships are found and prediction equations are based thereon, several precautions must be taken. First of all, no amount of mathematical manipulation can take into account all the weather

factors affecting the final yield. Hence any results obtained from the prediction equations must be carefully considered and readjusted if some weather factors not included in the equation definitely appear to be important in any particular instance. Furthermore, adjustments must be made in the light of any known significant changes in nonmeteorological factors.

APPENDIX A

TABLES

[p.38 blank]

Table 1

Simple Correlations between Spring Wheat Yields and Weather Factors in the USSR for Various Months by Uyezds [a]

Guberniya and Uyezd [b]	Precipitation								Average Maximum Temperature		
	May-Jun	May	Jun	Jul	Apr-Jun	Mar-Jun	Feb-Jun	Sep-Oct, May-Jun	May	Jun	Jul
Don											
Rostov	.126 (23)		.112 (23)				.241 (23)			-.024 (23)	
Ust'-Medveditskaya	.718**(17)						.575* (17)		-.395 (15)	-.550* (15)	-.545* (15)
Khoper	.553**(24)	.437* (24)	.473* (24)		.569**(24)	.643**(24)	.716**(24)		-.278 (24)	-.532**(24)	-.455* (24)
Saratov											
Tsaritsyn	.572* (15)								-.266 (15)	-.807 (15)	-.275 (15)
Kamyshin	.622**(23)									-.625**(23)	
Balashov	.506* (18)									-.578* (18)	
Atkarsk	.383 (26)								-.270 (25)	-.580**(26)	-.481* (26)
Saratov	.374 (25)									-.593**(25)	
Serdobsk	-.082 (18)									-.251 (18)	
Vol'sk	.410* (27)									-.568**(26)	
Samara											
Novouzensk	.488* (26)									-.398* (26)	
Nikolayevsk	.585* (18)									-.352 (18)	
Samara	.647**(27)				.582**(27)	.537**(27)	.529**(27)	.547**(27)	-.370 (27)	-.722**(27)	-.516**(27)
Kazan											
Kazan	.433* (28)									-.579**(28)	-.444* (28)
Koz'modem'yansk	.441* (22)								-.187 (24)	-.605**(24)	-.207 (24)

a. One asterisk in the table indicates a 5-percent level of significance; two asterisks indicate a 1-percent level of significance. Figures in parentheses indicate the number of observations. The minus sign (-) preceding a figure indicates an inverse relationship between the weather factor and the yield.
b. Guberniya corresponds roughly to present-day oblast; uyezd can be compared with present-day rayon.

[p 40 blank]

Table 1 (continued)

Simple Correlations between Spring Wheat Yields and Weather Factors in the USSR for Various Months by Uyezds

Guberniya and Uyezd	Precipitation								Average Maximum Temperature		
	May-Jun	May	Jun	Jul	Apr-Jun	Mar-Jun	Feb-Jun	Sep-Oct, May-Jun	May	Jun	Jul
Perm											
Krasnoufimsk	-.027 (18)									-.260 (18)	-.177 (18)
Yekaterinburg	.596**(24)									-.466* (24)	-.452* (24)
Kungur	.096 (26)									-.276 (26)	-.022 (26)
Irbit	.574* (18)									-.325 (18)	-.289 (18)
Perm	.324 (27)		.422* (27)	.135 (27)					.330 (27)	-.399* (27)	-.262 (27)
Solikamsk	.081 (16)									-.297 (16)	-.033 (15)
Cherdyn	.172 (22)		.248 (22)	-.053 (22)						-.248 (22)	-.037 (22)
Yekaterinoslav											
Yekaterinoslav	.363 (20)									-.222 (20)	
Aleksandrovsk	.293 (16)									.392 (16)	
Pavlograd	.540* (20)									.022 (20)	
Bakhmut	-.075 (17)									-.062 (17)	
Slavyanoserbsk	.353 (28)									-.379* (27)	
Voronezh											
Ostrogozhsk	.522* (20)									-.278 (19)	
Novokhopersk	.797**(14)									-.624* (14)	
Voronezh	.180 (21)	.093 (21)	.003 (21)						-.697**(21)	-.345 (21)	-.113 (21)
Zemlyansk	.208 (19)									-.380 (19)	
Zadonsk	.433 (19)									-.201 (19)	

Table 1 (continued)

Simple Correlations between Spr ing Wheat Yields and Weather Factors in the USSR for Various Months by Uyezds

Guberniya and Uyezd	Precipitation								Average Maximum Temperature		
	May-Jun	May	Jun	Jul	Apr-Jun	Mar-Jun	Feb-Jun	Sep-Oct, May-Jun	May	Jun	Jul
Tambov											
Tambov	.314 (24)									-.196 (24)	
Lipetsk	.187 (15)									-.278 (14)	
Kozlov	.049 (27)	.121 (27)	-.026 (27)							-.030 (27)	
Morshansk	-.070 (27)								-.493**(27)	-.214 (27)	-.311 (27)
Yelatomsk	.378 (20)									-.594**(20)	
Bessarabia											
Kishinev	.082 (25)								-.494* (25)	-.174 (25)	-.010 (25)
Soroki	-.081 (19)									-.399 (19)	
Kiev											
Uman	.159 (26)								-.286 (26)	-.343 (26)	-.094 (26)
Cherkassy	.401 (20)									-.469* (20)	
Kiev	.037 (26)	-.039 (26)	.081 (26)							-.126 (26)	
Radomysl	.651**(21)									-.366 (21)	

Table 2

Simple Correlations between Spring Barley Yields and Weather Factors in the USSR for Various Months by Uyezds a/

Guberniya and Uyezd b/	Precipitation						Average Maximum Temperature		
	May-Jun	May	Jun	Jul	Apr-Jun	Feb-Jun	May	Jun	Jul
Don									
Rostov	.050 (23)	.132 (23)							-.258 (23)
Ust'-Medveditskaya	.842**(17)	.347 (17)					-.204 (15)	-.418 (15)	-.666**(15)
Khoper	.513**(24)	.599**(24)	.276 (24)		.461* (24)	.435* (24)	-.331 (24)	-.294 (24)	-.452* (24)
Saratov									
Tsaritsyn	.002 (15)						-.423 (14)	-.533* (14)	.166 (15)
Kamyshin	.559**(22)							-.610**(22)	-.808**(22)
Balashov	.632**(17)							-.682**(17)	-.814**(17)
Atkarsk	.412* (25)	.391 (24)					-.474* (24)	-.547**(25)	-.563**(25)
Saratov	.483* (24)							-.584**(24)	-.438* (24)
Serdobsk	.178 (14)							-.243 (14)	-.457 (14)
Vol'sk	.651**(26)							-.619**(26)	-.628**(26)
Samara									
Novouzensk	.522**(26)							-.416* (26)	-.540**(26)
Nikolayevsk	.743**(18)							-.642**(18)	-.270 (18)
Samara	.757**(27)	.416* (27)	.570**(27)				-.371 (27)	-.710**(27)	-.581**(27)
Kazan									
Kazan	.318 (28)							-.597**(28)	-.409* (28)
Koz'modem'yansk	.238 (22)						-.177 (24)	-.506* (24)	-.189 (24)

a. One asterisk in the table indicates a 5-percent level of significance; two asterisks indicate a 1-percent level of significance. Figures in parentheses indicate the number of observations. The minus sign (-) preceding a figure indicates an inverse relationship between the weather factor and the yield.
b. Guberniya corresponds roughly to present-day oblast; uyezd can be compared with present-day rayon.

[p. 44 blank]

Table 2 (continued)

Simple Correlations between Spring Barley Yields and Weather Factors in the USSR for Various Months by Uyezds

Guberniya and Uyezd	Precipitation						Average Maximum Temperature		
	May-Jun	May	Jun	Jul	Apr-Jun	Feb-Jun	May	Jun	Jul
Perm									
Krasnoufimsk	.263 (18)							-.387 (18)	-.349 (18)
Yekaterinburg	.596**(24)							-.322 (24)	-.512* (24)
Kungur	.164 (26)							-.335 (26)	-.251 (26)
Irbit	.574* (18)							-.290 (18)	-.356 (18)
Perm	.350 (27)		.419* (27)	.158 (27)			.331 (27)	-.425* (27)	-.446* (27)
Solikamsk	.108 (16)							-.368 (16)	-.375 (15)
Cherdyn	.136 (22)		.141 (22)	.130 (22)				-.507* (22)	-.387 (22)
Yekaterinoslav									
Yekaterinoslav	.384 (20)							-.528* (20)	
Aleksandrovsk	.112 (16)							.249 (16)	
Pavlograd	.682**(20)							-.069 (20)	
Bakhmut	.126 (17)							-.102 (17)	
Slavyanoserbsk	.407* (28)							-.384* (27)	
Voronezh									
Ostrogozhsk	.487* (19)							-.356 (18)	-.349 (17)
Novokhopersk	.528* (14)							-.443 (14)	-.273 (14)
Voronezh	.444* (20)	.178 (20)	.211 (20)				-.446* (20)	-.303 (20)	-.035 (20)
Zemlyansk	.225 (18)						-.561* (18)	-.551* (18)	-.276 (18)
Zadonsk	.598* (14)							-.244 (14)	-.567* (14)

Table 2 (continued)

Simple Correlations between Spring Barley Yields and Weather Factors in the USSR for Various Months by Uyezds

Guberniya and Uyezd	Precipitation						Average Maximum Temperature		
	May-Jun	May	Jun	Jul	Apr-Jun	Feb-Jun	May	Jun	Jul
Tambov									
Tambov	.432* (22)							-.065 (22)	-.150 (22)
Lipetsk	.280 (13)							-.521 (13)	-.259 (13)
Kozlov	.220 (25)	.269 (25)	.038 (25)				.398* (25)	.021 (25)	-.124 (25)
Morshansk	.098 (22)						-.180 (22)	.133 (22)	.281 (22)
Yelatomsk	.334 (21)							-.298 (21)	-.297 (21)
Bessarabia									
Kishinev	.311 (25)						-.185 (25)	-.237 (25)	-.052 (25)
Soroki	.431 (19)							-.373 (19)	
Kiev									
Uman	.186 (26)						-.307 (26)	-.258 (26)	-.233 (26)
Cherkassy	.517* (20)							-.419 (20)	-.147 (20)
Kiev	.208 (28)	-.058 (28)	.319 (28)					-.367 (28)	-.425* (28)
Radomysl	.337 (22)						-.510* (21)	-.736**(22)	-.369 (22)

~~SECRET~~

Table 3

Simple Correlations between Oat Yields and Weather Factors in the USSR for Various Months by Uyezds [a]

Guberniya and Uyezd [b]	Precipitation								Average Maximum Temperature		
	May-Jun	May	Jun	Jul	Apr-May	Apr-Jun	Mar-Jun	Feb-Jun	May	Jun	Jul
Don											
Rostov	.274 (23)							.101 (23)			-.244 (23)
Ust'-Medveditskaya	.745**(17)							.552* (17)	-.154 (15)	-.513* (15)	-.450 (15)
Khoper	.583**(24)	.468* (24)	.439* (24)		.492* (24)	.603**(24)	.644**(24)	.737**(24)	-.355 (24)	-.521**(24)	-.519**(24)
Saratov											
Tsaritsyn	.188 (15)								-.405 (15)	-.760**(15)	-.365 (15)
Kamyshin	.579**(23)									-.673**(23)	-.542**(23)
Balashov	.687**(18)									-.717**(18)	-.689**(18)
Atkarsk	.361 (26)								-.292 (25)	-.661**(26)	-.689**(26)
Saratov	.497* (25)									-.644**(25)	-.649**(25)
Serdobsk	.223 (18)									-.454 (18)	-.597**(18)
Vol'sk	.430* (27)									-.566**(27)	-.527**(27)
Samara											
Novouzensk	.469* (26)									-.350 (26)	-.460* (26)
Nikolayevsk	.753**(18)									-.444 (18)	-.356 (18)
Samara	.656**(27)								-.363 (27)	-.726**(27)	-.630**(27)
Kazan											
Kazan	.334 (28)									-.431* (28)	-.383* (28)
Koz'modem'yansk	.116 (22)								-.247 (24)	-.393 (24)	-.156 (24)

a. One asterisk in the table indicates a 5-percent level of significance; two asterisks indicate a 1-percent level of significance. Figures in parentheses indicate the number of observations. The minus sign (-) preceding a figure indicates an inverse relationship between the weather factor and the yield.

b. Guberniya corresponds roughly to present-day oblast; uyezd can be compared with present-day rayon.

~~SECRET~~

[p.48 blank]

Table 3 (continued)

Simple Correlations between Oat Yields and Weather Factors in the USSR for Various Months by Uyezds

Guberniya and Uyezd	Precipitation								Average Maximum Temperature		
	May-Jun	May	Jun	Jul	Apr-May	Apr-Jun	Mar-Jun	Feb-Jun	May	Jun	Jul
Perm											
Krasnoufimsk	.046 (18)									-.442 (18)	-.351 (18)
Yekaterinburg	.533**(24)									-.291 (24)	-.392 (24)
Kungur	.322 (26)									-.296 (26)	-.270 (26)
Irbit	.667**(18)									-.247 (18)	-.316 (18)
Perm	.464* (27)		.008 (22)	.170 (27)					.187 (27)	-.574**(27)	-.553**(27)
Solikamsk	.289 (16)									-.431 (16)	-.398 (15)
Cherdyn	.139 (22)		.507**(27)	.261 (22)						-.381 (22)	-.349 (22)
Yekaterinoslav											
Yekaterinoslav	.280 (20)									-.325 (20)	
Aleksandrovsk	.367 (16)									.353 (16)	
Pavlograd	.431 (20)									-.076 (20)	
Bakhmut	.217 (16)									-.195 (16)	
Slavyanoserbsk	.406* (28)									-.398* (27)	
Voronezh											
Ostrogozhsk	.477**(20)									-.476* (19)	-.368 (18)
Novokhopersk	.728**(14)									-.635* (14)	-.532* (14)
Voronezh	.355 (22)	.256 (22)	-.295 (22)						-.726**(22)	-.361 (22)	-.301 (22)
Zemlyansk	.504* (19)									-.445 (19)	-.527* (19)
Zadonsk	.574**(19)									-.393 (19)	-.479* (19)

Table 3 (continued)

Simple Correlations between Oat Yields and Weather Factors in the USSR for Various Months by Uyezds

Guberniya and Uyezd	Precipitation								Average Maximum Temperature		
	May-Jun	May	Jun	Jul	Apr-May	Apr-Jun	Mar-Jun	Feb-Jun	May	Jun	Jul
Tambov											
Tambov	.556**(24)									-.556**(24)	-.419* (24)
Lipetsk	.478 (17)									-.640**(16)	-.387 (16)
Kozlov	.354 (27)	.335 (27)	.155 (27)							-.414* (27)	-.205 (27)
Morshansk	.463* (27)								-.448* (27)	-.450* (27)	-.484**(27)
Yelatomsk	.543**(22)									-.888**(22)	-.287 (22)
Bessarabia											
Kishinev	.385 (25)								-.045 (25)	-.254 (25)	-.173 (25)
Soroki	.309 (19)									-.254 (19)	
Kiev											
Uman	-.207 (26)								-.201 (26)	.143 (26)	-.059 (26)
Cherkassy	.562**(20)									-.588**(20)	-.208 (20)
Kiev	.369 (28)	-.095 (28)	.575**(28)							-.375* (28)	-.465* (28)
Radomysl	.406 (22)								-.616**(21)	-.452* (22)	-.545**(22)

Table 4

Simple Correlations, under the Given Hypothesis [a], between Spring Grain Yields and Weather Factors in the USSR for Various Months by Guberniyas [b]

Guberniya [c]	Number of Observations	Spring Wheat: Precipitation May-Jun	Spring Wheat: Average Maximum Temperature May	Spring Wheat: Average Maximum Temperature Jun	Spring Barley: Precipitation May-Jun	Spring Barley: Average Maximum Temperature May	Spring Barley: Average Maximum Temperature Jun	Oats: Precipitation May-Jun	Oats: Average Maximum Temperature May	Oats: Average Maximum Temperature Jun
Don	26	.387*		-.530**	.506**		-.536**	.560**		-.691**
Saratov	28	.681**		-.646**	.702**		-.228	.588**		-.672**
Samara	28	.719**		-.560**	.819**		-.613**	.657**		-.650**
Simbirsk	21	.421		-.619**	.569**		-.619**	.527*		-.648**
Penza	22	.430*		-.521*	.337		-.369	.497*		-.626**
Tambov	28	.292	-.554**		.573**	-.642**		.547**	-.564**	
Ryazan	25	.465*	-.391		.507**	-.411*		.676**	-.635**	
Tula	19	.431	-.562*		.155	-.130		.436	-.560*	
Orel	21	.177	-.353		.333	-.135		.027	-.372	
Kursk	21	.429	-.664**		.430	-.736**		.331	-.499*	
Voronezh	24	.663**	-.462*		.661**	-.444*		.586**	-.418*	
Chernigov	22	.394		-.750**	.593**		-.670**	.647**		-.704**
Poltava	24	.536**		-.470*	.579**		-.510*	.560**		-.560**
Bessarabia	22	.244		-.168	.396		-.253	.377		-.313
Yekaterinoslav	26	.191		-.145	.369		-.254	.413*		-.284

a. Sum of May and June precipitation hypothesized as "critical" for all guberniyas; June average maximum temperatures hypothesized as "critical" for all guberniyas except the six indicated in the table where May average maximum temperatures were considered "critical".
b. One asterisk in the table indicates a 5-percent level of significance; two asterisks indicate a 1-percent level of significance. The minus sign (-) preceding a figure indicates an inverse relationship between the weather factor and the yield.
c. Guberniya corresponds roughly to present-day oblast.

[p. 50 blank]

Table 5

Simple Correlations between Spring Grain Yields and Average Maximum Temperature in the USSR for Various Months by Guberniyas a/

Guberniya b/	Number of Observations	Spring Wheat May	Spring Wheat Jun	Spring Wheat Jul	Spring Barley May	Spring Barley Jun	Spring Barley Jul	Oats May	Oats Jun	Oats Jul
Don	26	-.442*		-.379	-.524**		-.377	-.506**		-.408*
Saratov	28	-.382*		-.484**	-.399*		-.498**	-.368		-.556**
Samara	28	-.412*		-.530**	-.416*		-.616**	-.400*		-.614**
Simbirsk	21	-.392		-.357	-.421		-.330	-.397		-.442*
Penza	22	-.509*		-.438*	-.560**		-.099	-.569**		-.514*
Tambov	28		-.198	-.055		-.320	-.224		-.588**	-.411*
Ryazan	25		-.482*	-.299		-.449*	-.510**		-.678**	-.420*
Tula	19		-.220	-.211		-.320	-.391		-.671**	-.545*
Orel	21		-.555**	-.394		-.398	-.703**		-.527*	-.581**
Kursk	21		-.462*	-.307		-.433*	-.448*		-.603**	-.499*
Voronezh	24		-.467*	-.223		-.490*	-.353		-.526**	-.292
Chernigov	22	-.484*		-.318	-.628**		-.476	-.454*		-.551
Poltava	24	-.461*		-.169	-.479*		-.412*	-.345		-.421*
Bessarabia	22	-.341		.144	-.198		.020	-.217		-.151
Yekaterinoslav	26	-.254		-.363	-.348		-.440*	-.296		-.410*

a. One asterisk in the table indicates a 5-percent level of significance; two asterisks indicate a 1-percent level of significance. The minus sign (-) preceding a figure indicates an inverse relationship between the weather factor and the yield.
b. Guberniya corresponds roughly to present-day oblast.

EB S4 K blank

Table 6

Multiple Correlations, under the Given Hypothesis a/, between Spring Grain Yields and Weather Factors in the USSR for Various Months by Guberniyas b/

Guberniya c/	Number of Observations	Spring Wheat: May-June Precipitation / Average Maximum Temperature — May	Spring Wheat: Jun	Spring Barley: May-June Precipitation / Average Maximum Temperature — May	Spring Barley: Jun	Oats: May-June Precipitation / Average Maximum Temperature — May	Oats: Jun
Don	26		.591**		.659**		.798**
Saratov	28		.795**		.704**		.758**
Samara	28		.753**		.849**		.752**
Simbirsk	21		.627*		.678**		.681**
Penza	22		.531*		.387		.633**
Tambov	28	.559**		.727**		.662**	
Ryazan	25	.491*		.531*		.745**	
Tula	19	.592*		.167		.592*	
Orel	21	.394		.334		.406	
Kursk	21	.701**		.762**		.530*	
Voronezh	24	.717**		.708**		.638**	
Chernigov	22		.760**		.764**		.815**
Poltava	24		.597**		.646**		.662**

a. Sum of May and June precipitation hypothesized as "critical" for all guberniyas; June average maximum temperatures hypothesized as "critical" for all guberniyas except the six indicated in the table where May average maximum temperatures were considered "critical".
b. One asterisk in the table indicates a 5-percent level of significance; two asterisks indicate a 1-percent level of significance.
c. Guberniya corresponds roughly to present-day oblast.

[p. 56 blank]

~~SECRET~~

Table 7

Multiple Correlations between Spring Grain Yields and Weather Factors in the USSR for Various Months by Guberniyas a/

Guberniya b/	Number of Observations	Spring Wheat: May-June Precipitation, Average Maximum Temperature — May	Jun	Jul	Spring Barley: May-June Precipitation, Average Maximum Temperature — May	Jun	Jul	Oats: May-June Precipitation, Average Maximum Temperature — May	Jun	Jul
Don	26	.528*		.496*	.654**		.580**	.677**		.638**
Saratov	28	.700**		.721**	.723**		.726**	.640**		.667**
Samara	28	.734**		.729**	.827**		.834**	.677**		.712**
Simbirsk	21	.499		.464	.620*		.577*	.578*		.579*
Penza	22	.552*		.513	.566*		.341	.625**		.599*
Tambov	28		.309	.302		.588**	.575**		.692**	.602**
Ryazan	25		.535*	.481		.544*	.609**		.765**	.696**
Tula	19		.448	.431		.330	.392		.729**	.578*
Orel	21		.568*	.395		.494	.705**		.528	.623*
Kursk	21		.575*	.468		.556*	.544*		.639**	.534*
Voronezh	24		.723**	.667**		.731**	.694**		.695**	.609**
Chernigov	22	.591*		.458	.817**		.688**	.750**		.767**
Poltava	24	.629**		.541*	.670**		.657**	.596**		.646**

a. One asterisk in the table indicates a 5-percent level of significance; two asterisks indicate a 1-percent level of significance.
b. Guberniya corresponds roughly to present-day oblast.

~~SECRET~~

[p. 58 blank]

Table 8

Multiple Regression Equations, Based on May-June Precipitation (X_1) and Average Maximum June Temperature (X_2), For Use in Predicting Spring Grain Yields in Selected Guberniyas a/ of the USSR

Guberniya	Number of Observations	Spring Wheat	Spring Barley	Oats
Don	26	$77.4 + 0.123X_1 - 2.183X_2$	$93.5 + 0.243X_1 - 2.798X_2$	$104.0 + 0.244X_1 - 3.624X_2$
Saratov	28	$68.5 + 0.211X_1 - 2.351X_2$	$-20.4 + 0.531X_1 + 0.534X_2$	$118.1 + 0.249X_1 - 4.264X_2$
Samara	28	$55.9 + 0.325X_1 - 2.049X_2$	$54.2 + 0.455X_1 - 2.464X_2$	$98.3 + 0.270X_1 - 3.689X_2$
Simbirsk	21	$118.8 + 0.056X_1 - 4.257X_2$	$134.5 + 0.240X_1 - 5.176X_2$	$135.4 + 0.158X_1 - 5.303X_2$
Penza	22	$89.2 + 0.091X_1 - 2.768X_2$	$61.7 + 0.145X_1 - 1.692X_2$	$112.6 + 0.093X_1 - 3.844X_2$
Tambov	28	$52.0 + 0.186X_1 - 0.809X_2$	$33.9 + 0.403X_1 - 1.070X_2$	$98.6 + 0.295X_1 - 3.472X_2$
Ryazan	25	$80.1 + 0.231X_1 - 2.438X_2$	$54.9 + 0.323X_1 - 1.908X_2$	$76.1 + 0.291X_1 - 2.708X_2$
Tula	19	$49.7 + 0.171X_1 - 0.657X_2$	$78.4 + 0.043X_1 - 1.907X_2$	$89.6 + 0.101X_1 - 2.452X_2$
Orel	21	$119.7 + 0.059X_1 - 3.507X_2$	$59.2 + 0.110X_1 - 1.789X_2$	$91.7 - 0.009X_1 - 2.230X_2$
Kursk	21	$75.8 + 0.121X_1 - 2.011X_2$	$79.2 + 0.145X_1 - 2.187X_2$	$93.7 + 0.068X_1 - 2.609X_2$
Voronezh	24	$57.2 + 0.225X_1 - 1.933X_2$	$65.0 + 0.254X_1 - 2.415X_2$	$59.4 + 0.083X_1 - 1.128X_2$
Chernigov	22	$112.0 + 0.036X_1 - 2.776X_2$	$93.5 + 0.142X_1 - 2.590X_2$	$74.2 + 0.124X_1 - 2.089X_2$
Poltava	24	$73.6 + 0.158X_1 - 1.678X_2$	$78.9 + 0.176X_1 - 1.880X_2$	$91.1 + 0.142X_1 - 2.099X_2$

a. Guberniya corresponds roughly to present-day oblast.

[p.60 blank]

APPENDIX B

SOURCES

1. Urozhay (Annual Publications of the Central Statistical Committee), Ministry of the Interior, St. Petersburg, 1883-1915.
2. Letopici (Annual Publications of the Central Physical Observatory), St. Petersburg, 1883-1915.
3. Vladimir P. Timoshenko, Agricultural Russia and the Wheat Problem, Stanford University Food Research Institute, 1932, p. 279.
4. Naum Jasny, The Socialized Agriculture of the USSR, Stanford University Food Research Institute, 1949, p. 497.
5. George W. Snedecor, Statistical Methods, Iowa State College Press, 1946, pp. 149 and 351.